

Completely positive and completely positive semidefinite tensor relaxations for polynomial optimization

Xiaolong Kuang¹ · Luis F. Zuluaga¹

Received: 7 November 2016 / Accepted: 17 August 2017 / Published online: 30 August 2017
© Springer Science+Business Media, LLC 2017

Abstract Completely positive (CP) tensors, which correspond to a generalization of CP matrices, allow to reformulate or approximate a general polynomial optimization problem (POP) with a conic optimization problem over the cone of CP tensors. Similarly, completely positive semidefinite (CPSD) tensors, which correspond to a generalization of positive semidefinite (PSD) matrices, can be used to approximate general POPs with a conic optimization problem over the cone of CPSD tensors. In this paper, we study CP and CPSD tensor relaxations for general POPs and compare them with the bounds obtained via a Lagrangian relaxation of the POPs. This shows that existing results in this direction for quadratic POPs extend to general POPs. Also, we provide some tractable approximation strategies for CP and CPSD tensor relaxations. These approximation strategies show that, with a similar computational effort, bounds obtained from them for general POPs can be tighter than bounds for these problems obtained by reformulating the POP as a quadratic POP, which subsequently can be approximated using CP and PSD matrices. To illustrate our results, we numerically compare the bounds obtained from these relaxation approaches on small scale fourth-order degree POPs.

Keywords Copositive programming · Convex relaxation · Completely positive tensor · Completely positive semidefinite tensor

1 Introduction

Polynomials appear in a wide variety of areas in science. It is not surprising then that optimizing a polynomial optimization problem (POP), in which both the objective function

✉ Xiaolong Kuang
kuangxiaolong0731@gmail.com

Luis F. Zuluaga
luis.zuluaga@lehigh.edu

¹ Industrial and Systems Engineering, Lehigh University, Bethlehem, PA 18015, USA

and constraints are defined by multivariate polynomials, continues to be the focus of novel research work (cf., [3]). Here, the interest is the class of non-convex, non-linear POPs. Clearly, a non-convex quadratic program (QP) belongs to this class of problems, and its study has been widely addressed in the literature. For example, semidefinite programming (SDP) relaxations have been actively used to find good bounds and approximate solutions for general QPs (see, e.g., [17,36,50]), and for important QP problems such as the max-cut and the stable set problem (see, e.g., [18,19,24,41]). In [27], more computationally efficient second order cone programming (SOCP) relaxations have also been proposed to approximately solve non-convex QPs.

The early work linking convex optimization and polynomial optimization in [39,47] revealed the possibility to use conic optimization to obtain global or near-global solutions for non-convex POPs in which polynomials with degree higher than two are used. In the seminal work of Parrilo [40] and Lasserre [29], SDP is used to obtain the global or near-global optimum of POPs. Besides SDP approximations, other convex approximations to address the solution of POPs have been investigated using linear programming (LP) and SOCP techniques [1,30,31,42,52]. These techniques are at the core of the well-known area of polynomial optimization (cf., [3]).

Alternatively, it has been shown that several NP-hard optimization problems can be reformulated as a *completely positive (CP) program*; that is, a linear program over the convex cone of CP matrices or its dual cone, the cone of copositive matrices, including standard QPs [10], stable set problems [18,22], graph partitioning problems [44], and quadratic assignment problems [45]. In [13], Burer derives a more general result; namely, that every linearly constrained QP with binary variables can be reformulated as a copositive program. CP programming relaxations for general quadratically constrained quadratic programs (QCQPs) have been studied in [4,15]. In [6], CP programming reformulation for QCQPs and QPs with complementarity constraints (QPCCs) are discussed without any boundedness assumption on the problems' feasible regions. Although the CP matrix cone is not tractable in general, recent advances on obtaining approximation algorithms [2,12,20] for CP programs, provide an alternative way to globally solve QCQPs. Recently, Bomze shows in [8] that CP programming relaxations provide tighter bounds than Lagrangian dual bounds for quadratically and linearly constrained QPs.

A natural thought is whether one can extend the copositive programming or completely positive programming reformulations for QPs to POPs. Arima et al. [5] propose the moment cone relaxation for a class of POPs to extend the results on CP programming relaxations for the QCQPs. Recently, Peña et al. [43] show that under certain conditions, general POPs can be reformulated as a conic program over the cone of CP tensors, which is a natural extension of the cone of CP matrices used for the solution of quadratic POPs. This tensor representation was originally proposed in [21], and is now the focus of active research (see, e.g., [25,26,35,48]). In [43], it is also shown that the conditions for the equivalence between POPs and its associated CP programming relaxation, when applied to QCQPs, lead to conditions that are weaker than the ones introduced in [13].

In this article, we study CP and *completely positive semidefinite (CPSD)* tensor relaxations for POPs (cf., Sect. 2.1). Our main contributions are: (1) We extend the results for QPs in [8] to general POPs by using CP and CPSD tensor cones. In particular, we show that CP tensor relaxations provide tighter bounds than Lagrangian relaxations for general POPs. (2) We provide tractable approximations for CP and CPSD tensor cones that can be used to globally approximate general POPs. (3) We prove that CP tensor relaxations yield tighter bounds than CP matrix relaxations for quadratic reformulations of some classes of POPs. (4) We provide numerical results to show that, in more generality, approximations to the CP

tensor relaxations of a POP can be used to obtain a tighter bound for the problem than the one obtained based on approximations to the CP matrix relaxation of the POP’s associated quadratic reformulation.

The remainder of the article is organized as follows. We briefly introduce the basic concepts of tensor cones and tensor representations of polynomials in Sect. 2. Lagrangian relaxation, CP and CPSD tensor relaxations for POPs are discussed in Sect. 3. In Sect. 4, we discuss the *quadratic reformulation* of a general POP; that is, auxiliary decision variables are introduced to the problem to reformulate it as a QCQP. Then, for a class of POPs, the bounds on the POP obtained from a CP matrix relaxations of the quadratic reformulation of the POP are compared with the ones obtained from a CP tensor relaxation of the POP. In Sect. 5, linear matrix inequality (LMI) approximation strategies for the CP and CPSD tensor cones are proposed, and a comparison of bounds obtained by tensor relaxation and matrix relaxation is presented for general small scale POPs. Lastly, Sect. 6 summarizes the article’s results and provides future working directions.

2 Preliminaries

2.1 Basic concepts and notation

We first introduce basic concepts and the notation used throughout the article. Following [43], we start by defining tensors.

Definition 1 Let $\mathcal{T}_{n,d}$ denote the set of *tensors* of dimension n and order d in \mathbb{R}^n , that is

$$\mathcal{T}_{n,d} = \underbrace{\mathbb{R}^n \otimes \cdots \otimes \mathbb{R}^n}_d,$$

where \otimes is the tensor product.

A tensor $T \in \mathcal{T}_{n,d}$ is symmetric if the entries are invariant with respect to permutations of its indices. We denote by $\mathcal{S}_{n,d} \subseteq \mathcal{T}_{n,d}$ the set of symmetric tensors of dimension n and order d . For any $T^1, T^2 \in \mathcal{T}_{n,d}$, let $\langle \cdot, \cdot \rangle_{n,d}$ denote the tensor inner product defined by

$$\langle T^1, T^2 \rangle_{n,d} = \sum_{\{i_1, \dots, i_d\} \in \{1, \dots, n\}^d} T^1_{(i_1, \dots, i_d)} T^2_{(i_1, \dots, i_d)}.$$

Definition 2 For any $x \in \mathbb{R}^n$, let the mapping $\mathbb{R}^n \rightarrow \mathcal{S}_{n,d}$ be defined by

$$M_d(x) = \underbrace{x \otimes \cdots \otimes x}_d.$$

Definitions 1 and 2 are natural extensions of matrix notation to tensors. For example, $\mathcal{T}_{n,2}$ is the set of $n \times n$ matrices, while $\mathcal{S}_{n,2}$ is the set of $n \times n$ symmetric matrices. Also, $\langle \cdot, \cdot \rangle_{n,2}$ is the Frobenius inner product and $M_2(x) = xx^\top$ for any $x \in \mathbb{R}^n$. In general, $M_d(x)$ is the symmetric tensor whose (i_1, \dots, i_d) entry is $x_{i_1} \cdots x_{i_d}$.

Proposition 1 Let $\mathbb{E}_{n,d}$ be all-ones tensor with dimension n and order d and $e \in \mathbb{R}^n$ be the all-ones vector, then

$$\langle \mathbb{E}_{n,d}, M_d(x) \rangle_{n,d} = (e^\top x)^d, \forall x \in \mathbb{R}^n.$$

Proof By the definition of $M_d(\cdot)$ and $\langle \cdot, \cdot \rangle_{n,d}$,

$$\langle \mathbb{E}_{n,d}, M_d(x) \rangle_{n,d} = \sum_{k_1+k_2+\dots+k_n=d} \binom{d}{k_1, k_2, \dots, k_n} x_1^{k_1} x_2^{k_2} \dots x_n^{k_n} = (e^\top x)^d,$$

where $\binom{d}{k_1, k_2, \dots, k_n}$ is the multinomial coefficient. □

Proposition 2 For $x \in \mathbb{R}^n, y \in \mathbb{R}^n$,

$$\langle M_d(x), M_d(y) \rangle_{n,d} = (x^\top y)^d.$$

Proof Let $x, y \in \mathbb{R}^n$ be given and $z \in \mathbb{R}^n$ be defined as $z_i = x_i y_i, i = 1, \dots, n$, and let $e \in \mathbb{R}^n$ be the all-ones vector. From the definition of $M_d(\cdot)$ and $\langle \cdot, \cdot \rangle_{n,d}$ we have

$$\begin{aligned} \langle M_d(x), M_d(y) \rangle_{n,d} &= \sum_{\{i_1, \dots, i_d\} \in \{1, \dots, n\}^d} M_d(x)_{(i_1, \dots, i_d)} M_d(y)_{(i_1, \dots, i_d)} \\ &= \sum_{\{i_1, \dots, i_d\} \in \{1, \dots, n\}^d} x_{i_1} x_{i_2} \dots x_{i_d} \cdot y_{i_1} y_{i_2} \dots y_{i_d} \\ &= \sum_{\{i_1, \dots, i_d\} \in \{1, \dots, n\}^d} (x_{i_1} y_{i_1}) (x_{i_2} y_{i_2}) \dots (x_{i_d} y_{i_d}) \\ &= \langle \mathbb{E}_{n,d}, M_d(z) \rangle_{n,d} \\ &= (e^\top z)^d \quad \text{(from Proposition 1)} \\ &= (x^\top y)^d. \end{aligned}$$

□

Analogous to PSD and copositive matrices, PSD and copositive tensors can be defined as follows.

Definition 3 Define the \mathcal{K} -semidefinite (or set-semidefinite) symmetric tensor cone of dimension n and order d as:

$$\mathcal{C}_{n,d}(\mathcal{K}) = \{T \in \mathcal{S}_{n,d} : \langle T, M_d(x) \rangle_{n,d} \geq 0, \forall x \in \mathcal{K}\}.$$

For $\mathcal{K} = \mathbb{R}^n, \mathcal{C}_{n,d}(\mathbb{R}^n)$ denotes the positive semidefinite (PSD) tensor cone. For $\mathcal{K} = \mathbb{R}_+^n, \mathcal{C}_{n,d}(\mathbb{R}_+^n)$ denotes the copositive tensor cone.

Similar to the one-to-one correspondence of $n \times n$ PSD matrices to nonnegative homogeneous quadratic polynomials of n variables, there is also a one-to-one correspondence of PSD tensors with dimension n and order d to nonnegative homogeneous polynomials with n variables and degree d (cf., [35]). Note that the set of non-negative homogeneous polynomials of odd degree is empty; that is, $\mathcal{C}_{n,d}(\mathbb{R}^n) = \emptyset$ when d is odd. Next we discuss the dual cones of $\mathcal{C}_{n,d}(\mathbb{R}_+^n)$ and $\mathcal{C}_{n,d}(\mathbb{R}^n)$, following the discussion in [35] and [43].

Definition 4 Given any cone $\mathcal{C} \subseteq \mathcal{S}_{n,d}$, the dual cone of \mathcal{C} is

$$\mathcal{C}^* = \{Y \in \mathcal{S}_{n,d} : \langle X, Y \rangle_{n,d} \geq 0, \forall X \in \mathcal{C}\}.$$

If $\mathcal{C}^* = \mathcal{C}$, then cone \mathcal{C} is self-dual.

The dual cones of the PSD tensor cone and copositive tensor cone have been studied in [35, 43]. Formally,

Proposition 3

- (a) $C_{n,d}^*(\mathbb{R}_+^n) = \text{conv}\{M_d(x) : x \in \mathbb{R}_+^n\}$.
- (b) $C_{n,2d}^*(\mathbb{R}^n) = \text{conv}\{M_{2d}(x) : x \in \mathbb{R}^n\}$.

Similar to the cone of CP and PSD matrices, we denote by $C_{n,d}^*(\mathbb{R}_+^n)$ the *completely positive (CP) tensor cone*, and by $C_{n,2d}^*(\mathbb{R}^n)$ the *completely positive semidefinite (CPSD) tensor cone*. Also, let the *homogeneous sum of square (SOS) tensor cone* of dimension n and order $2d$ be defined by

$$C_{n,2d}(\text{SOS}) = \left\{ T_{n,2d} : \langle T_{n,2d}, M_{2d}(x) \rangle_{n,2d} = \sum_{i \in \mathbb{N}} \lambda_i \left(\langle T_{n,d}^i, M_d(x) \rangle_{n,d} \right)^2, \lambda_i \geq 0, \forall i \in \mathbb{N} \right\}.$$

It is well known that the cone of PSD matrices is self-dual; however, in general, the PSD tensor cone is not self-dual (cf., [35]) as discussed next.

Proposition 4 ([35, Prop. 5.8 (i)])

$$C_{n,2d}^*(\mathbb{R}^n) \subseteq C_{n,2d}(\text{SOS}) \subseteq C_{n,2d}(\mathbb{R}^n).$$

Proof Let $T \in C_{n,2d}^*(\mathbb{R}^n)$, by Proposition 3, $T = \sum_{i \in \mathbb{N}} \lambda_i M_{2d}(y^i)$, where $y^i \in \mathbb{R}^n, \lambda_i \geq 0$, for all $i \in \mathbb{N}$ and $\sum_{i \in \mathbb{N}} \lambda_i = 1$. Then $\forall x \in \mathbb{R}^n$,

$$\begin{aligned} \langle T, M_{2d}(x) \rangle_{n,2d} &= \left\langle \sum_i \lambda_i M_{2d}(y^i), M_{2d}(x) \right\rangle_{n,2d} \\ &= \sum_{i \in \mathbb{N}} \lambda_i \langle M_{2d}(y^i), M_{2d}(x) \rangle_{n,2d} \\ &= \sum_{i \in \mathbb{N}} \lambda_i (x^\top y^i)^{2d} \quad (\text{from Proposition 2}) \\ &= \sum_{i \in \mathbb{N}} \left[\sqrt{\lambda_i} (x^\top y^i)^d \right]^2. \end{aligned}$$

Take $z_k^i = x_k y_k^i$, then $x^\top y^i = e^\top z^i$ where $e \in \mathbb{R}^n$ is the all-ones vector for all $i \in \mathbb{N}$. Therefore,

$$\begin{aligned} \langle T, M_{2d}(x) \rangle_{n,2d} &= \sum_{i \in \mathbb{N}} \left[\sqrt{\lambda_i} (e^\top z^i)^d \right]^2 \\ &= \sum_{i \in \mathbb{N}} \left[\sqrt{\lambda_i} \langle \mathbb{E}_{n,d}, M_d(z^i) \rangle_{n,d} \right]^2 \quad (\text{from Proposition 1}). \end{aligned}$$

Therefore, $C_{n,2d}^*(\mathbb{R}^n) \subseteq C_{n,2d}(\text{SOS})$. By the definition of the homogeneous SOS tensor cone, it is clear that $C_{n,2d}(\text{SOS}) \subseteq C_{n,2d}(\mathbb{R}^n)$. □

The proof of Proposition 4 can be seen as an alternative proof for Proposition 5.8 (i) in [35] that uses the tensor notation introduced in this article. As mentioned before, it is well known that $C_{n,2}^*(\mathbb{R}^n) = C_{n,2}(\text{SOS}) = C_{n,2}(\mathbb{R}^n)$. This statement coincides with the self-duality of the cone of PSD matrices. Luo et al. showed in [35] that $C_{n,2d}^*(\mathbb{R}^n) \subsetneq C_{n,2d}(\text{SOS})$ for $d \geq 2$. On the other hand, the Motzkin polynomial together with the isomorphism between homogeneous polynomials and tensors shows that $C_{n,2d}(\text{SOS}) \subsetneq C_{n,2d}(\mathbb{R}^n)$ when $d \geq 2$ and $n \geq 2$.

2.2 Tensor representation of general polynomials

In Sect. 2.1, we mentioned the isomorphism between symmetric tensors and homogenous polynomials. Next, we introduce a tensor representation for general polynomials that are not necessarily homogeneous. Define $\mathbb{R}[x]$ as the ring of polynomials with real coefficients in \mathbb{R}^n , and let $\mathbb{R}_d[x] := \{p \in \mathbb{R}[x] : \deg(p) \leq d\}$ denote the set of polynomials with dimension n and degree at most d . For simplicity, we use $M_d(1, x), x \in \mathbb{R}^n$ to represent $M_d((1, x^\top)^\top), x \in \mathbb{R}^n$ and use $T_d(p)$ to represent $T_d(p(x))$ for $p(x) \in \mathbb{R}_d[x]$ throughout the article. Then, for any $p(x) \in \mathbb{R}_d[x]$, we have

$$p(x) = \langle T_d(p), M_d(1, x) \rangle_{n+1,d}, \tag{1}$$

where $T_d(\cdot)$ is the mapping of coefficients of $p(x)$ in terms of $M_d(1, x)$ in $\mathcal{S}_{n+1,d}$. Following [43], define $T_d : \mathbb{R}_d[x] \rightarrow \mathcal{S}_{n+1,d}$ as

$$T_d \left(\sum_{\beta \in \mathbb{Z}_+^n : |\beta| \leq d} p_\beta x^\beta \right)_{i_1, \dots, i_d} := \frac{\alpha_1! \cdots \alpha_n!}{|\alpha|!} p_\alpha,$$

where α is the (unique) exponent such that $x^\alpha := x_1^{\alpha_1} \cdots x_n^{\alpha_n} = x_{i_1} \cdots x_{i_d}$ (i.e., α_k is the number of times k appears in the multi-set $\{i_1, \dots, i_d\}$) and $|\alpha| = \sum_{i=1}^n \alpha_i$. For any polynomial $p(x) \in \mathbb{R}_d[x]$, let $\tilde{p}(x)$ denote the homogenous component of $p(x)$ with the highest degree, then it follows

$$\tilde{p}(x) = \langle T_d(p), M_d(0, x) \rangle_{n+1,d}. \tag{2}$$

Equations (1) and (2) can be used to characterize the boundedness of general polynomials using their associated tensor representation.

Theorem 1 *Let $\mu \in \mathbb{R}$ and*

- (a) $p(x) \in \mathbb{R}_d[x]$. Then $p(x) \geq \mu$ for all $x \in \mathbb{R}_+^n$ if and only if $T_d(p - \mu) \in \mathcal{C}_{n+1,d}(\mathbb{R}_+^{n+1})$.
- (b) $p(x) \in \mathbb{R}_{2d}[x]$. Then $p(x) \geq \mu$ for all $x \in \mathbb{R}^n$ if and only if $T_{2d}(p - \mu) \in \mathcal{C}_{n+1,2d}(\mathbb{R}^{n+1})$.

Proof For (a), assume $T_d(p - \mu) \in \mathcal{C}_{n+1,d}(\mathbb{R}_+^{n+1})$. By Definition 3, $\langle T_d(p - \mu), M_d(1, x) \rangle_{n+1,d} \geq 0, \forall x \in \mathbb{R}_+^n$, then

$$p(x) - \mu = \langle T_d(p - \mu), M_d(1, x) \rangle_{n+1,d} \geq 0, \forall x \in \mathbb{R}_+^n. \tag{3}$$

For the other direction, assume $p(x) \geq \mu, \forall x \in \mathbb{R}_+^n$, then by (3), $\langle T_d(p - \mu), M_d(1, x) \rangle_{n+1,d} \geq 0, \forall x \in \mathbb{R}_+^n$. Thus, for any $(x_0, x) \in \mathbb{R}_{++} \times \mathbb{R}_+^n$,

$$\langle T_d(p - \mu), M_d(x_0, x) \rangle_{n+1,d} = x_0^d \left\langle T_d(p - \mu), M_d \left(1, \frac{x}{x_0} \right) \right\rangle_{n+1,d} \geq 0. \tag{4}$$

Furthermore, from the continuity of polynomials, we have that for $k > 0$,

$$\langle T_d(p - \mu), M_d(0, x) \rangle = \lim_{k \rightarrow +\infty} \langle T_d(p - \mu), M_d(1/k, x) \rangle \geq 0, \tag{5}$$

where the last inequality follows from (4). From (4), (5), and Definition 3, it follows that $T_d(p - \mu) \in \mathcal{C}_{n+1,d}(\mathbb{R}_+^{n+1})$.

The proof of (b) is similar to the proof of (a). □

Corollary 1 *Let $\mu \in \mathbb{R}$ and*

- (a) $p(x) \in \mathbb{R}_d[x]$. Then $\inf\{p(x) : x \in \mathbb{R}_+^n\} = \sup\{\mu \in \mathbb{R} : T_d(p - \mu) \in \mathcal{C}_{n+1,d}(\mathbb{R}_+^{n+1})\}$.

(b) $p(x) \in \mathbb{R}_{2d}[x]$. Then $\inf\{p(x) : x \in \mathbb{R}^n\} = \sup\{\mu \in \mathbb{R} : T_{2d}(p-\mu) \in \mathcal{C}_{n+1,2d}(\mathbb{R}^{n+1})\}$.

Theorem 1 and Corollary 1 generalize the key Lemma 2.1 and Corollary 2.1 in [8] for polynomials with degree higher than 2 by using a tensor representation. Moreover, Corollary 1 can be seen as a convexification of an unconstrained (possibly non-linear non-convex) POP to a linear conic program over the CP and CSDP tensor cones. In the next section, we will discuss CP and CPSD tensor relaxations for general constrained POPs.

3 Lagrangian and conic relaxations of POPs

Let $p_i(x) \in \mathbb{R}_d[x]$, $i = 0, \dots, m$. Consider two general POPs with polynomial constraints:

$$\begin{aligned} z_+ = \inf \quad & p_0(x) \\ \text{s.t.} \quad & p_i(x) \leq 0, \quad i = 1, \dots, m, \\ & x \in \mathbb{R}_+^n, \end{aligned} \tag{6}$$

and

$$\begin{aligned} z = \inf \quad & p_0(x) \\ \text{s.t.} \quad & p_i(x) \leq 0, \quad i = 1, \dots, m, \end{aligned} \tag{7}$$

where $d = \max\{\deg(p_i(x)) : i \in \{0, 1, \dots, m\}\}$ is the degree of the POP. Problems (6) and (7) represent general POPs, which encompass a large class of non-linear non-convex problems, including non-convex QPs with binary variables (i.e., binary constraints can be written in the polynomial form $x_i(1 - x_i) \leq 0$, $-x_i(1 - x_i) \leq 0$). Naturally, we have $z \leq z_+$ since the feasible set of problem (6) is a subset of problem (7). Next we show that the results of Bomze for QPs in [8] can be extended to POPs of form (6) and (7).

3.1 Lagrangian relaxations

Let $u_i \geq 0$ be the Lagrangian multiplier of the inequality constraints $p_i(x) \leq 0$ for $i = 1, \dots, m$ and $v_i \geq 0$ for constraints $x_i \in \mathbb{R}_+$ for $i = 1, \dots, n$. The Lagrangian function for problem (6) is

$$L_+(x; u, v) := p_0(x) + \sum_{i=1}^m u_i p_i(x) - v^\top x,$$

and the Lagrangian dual function of problem (6) is

$$\Theta_+(u, v) := \inf\{L_+(x; u, v) : x \in \mathbb{R}^n\},$$

with its optimal value

$$z_{LD,+} = \sup\{\Theta_+(u, v) : (u, v) \in \mathbb{R}_+^m \times \mathbb{R}_+^n\},$$

We also use a semi-Lagrangian dual function to represent the nonnegative variable constraints of problem (6),

$$\Theta_{semi}(u) := \inf\{L(x; u) : x \in \mathbb{R}_+^n\},$$

where $L(x; u) := p_0(x) + \sum_{i=1}^m u_i p_i(x)$, with its optimal value

$$z_{semi} = \sup\{\Theta_{semi}(u) : u \in \mathbb{R}_+^m\}.$$

Similarly, let $u_i \geq 0$ be the Lagrangian multiplier of the inequality constraints $p_i(x) \leq 0$ for $i = 1, \dots, m$. The Lagrangian function for problem (7) is

$$L(x; u) := p_0(x) + \sum_{i=1}^m u_i p_i(x),$$

and the Lagrangian dual function of problem (7) is

$$\Theta(u) := \inf\{L(x; u) : x \in \mathbb{R}^n\},$$

with its optimal value

$$z_{LD} = \sup\{\Theta(u) : u \in \mathbb{R}_+^m\}.$$

Thus we have the following relationship:

$$\begin{aligned} \Theta_+(u, v) &= \inf\{L_+(x; u, v) : x \in \mathbb{R}^n\} \\ &\leq \inf\{L_+(x; u, v) : x \in \mathbb{R}_+^n\} \\ &= \inf\{L(x; u) - v^T x : x \in \mathbb{R}_+^n\} \\ &\leq \inf\{L(x; u) : x \in \mathbb{R}_+^n\} = \Theta_{semi}(u), \end{aligned}$$

where the second inequality holds because $x, v \in \mathbb{R}_+^n$ always implies $v^T x \geq 0$. Therefore, we have:

$$z_{LD,+} \leq z_{semi} \leq z_+,$$

where the latter inequality holds by weak duality. Similarly, from weak duality theory, we have $z_{LD} \leq z$.

3.2 CPSD tensor relaxation for POP with free variables

Consider the following conic program:

$$\begin{aligned} z_{SP} &= \inf \quad \langle T_d(p_0), X \rangle \\ \text{s.t.} \quad &\langle T_d(p_i), X \rangle \leq 0, \quad i = 1, \dots, m, \\ &\langle T_d(1), X \rangle = 1, \\ &X \in \mathcal{C}_{n+1,d}^*(\mathbb{R}^{n+1}), \end{aligned} \tag{8}$$

and its conic dual problem

$$z_{SD} = \sup \left\{ \mu : T_d(p_0) - \mu T_d(1) + \sum_{i=1}^m u_i T_d(p_i) \in \mathcal{C}_{n+1,d}(\mathbb{R}^{n+1}), u \in \mathbb{R}_+^m \right\}. \tag{9}$$

Recall that $\mathcal{C}_{n+1,d}(\mathbb{R}^{n+1})$ is empty if d is odd. Thus, for simplicity, in what follows we assume without loss of generality that d is even in (8) (if this is not the case, one can change $d \rightarrow 2\lceil d/2 \rceil$ by adding explicit zeros to higher order monomials in $p_i, i = 0, 1, \dots, m$). Also we use $\langle \cdot, \cdot \rangle$ represent the tensor inner product of appropriate dimension and order.

Proposition 5 *Problem (8) is a relaxation of problem (7) with $z_{SP} \leq z$.*

Proof Let $x \in \mathbb{R}^n$ be a feasible solution of problem (7). It follows that $X = M_d(1, x)$ is a feasible solution of problem (8) directly by applying (1). Also $p(x) = \langle T_d(p_0), X \rangle$ is a direct result of (1) with the same objective value. □

Theorem 2 For problem (7), its Lagrangian dual function optimal value satisfies,

$$z_{LD} = \sup \{ \mu : (\mu, u) \in \mathbb{R} \times \mathbb{R}_+^m, T_d(L(x; u) - \mu) \in \mathcal{C}_{n+1,d}(\mathbb{R}^{n+1}) \}$$

and $z_{LD} = z_{SD} \leq z_{SP} \leq z$.

Proof By Corollary 1 (b),

$$\begin{aligned} \Theta(u) &= \inf \{ L(x; u) : x \in \mathbb{R}^n \} \\ &= \sup \{ \mu : T_d(L(x; u) - \mu) \in \mathcal{C}_{n+1,d}(\mathbb{R}^{n+1}) \}, \end{aligned}$$

then

$$\begin{aligned} z_{LD} &= \sup \{ \Theta(u) : u \in \mathbb{R}_+^m \} \\ &= \sup \{ \mu : (\mu, u) \in \mathbb{R} \times \mathbb{R}_+^m, T_d(L(x; u) - \mu) \in \mathcal{C}_{n+1,d}(\mathbb{R}^{n+1}) \}. \end{aligned}$$

From (9), we have

$$\begin{aligned} z_{SD} &= \sup \left\{ \mu : T_d(p_0) - \mu T_d(1) + \sum_{i=1}^m u_i T_d(p_i) \in \mathcal{C}_{n+1,d}(\mathbb{R}^{n+1}), u \in \mathbb{R}_+^m \right\} \\ &= \sup \left\{ \mu : T_d \left(p_0 + \sum_{i=1}^m u_i p_i - \mu \right) \in \mathcal{C}_{n+1,d}(\mathbb{R}^{n+1}), u \in \mathbb{R}_+^m \right\} \\ &= \sup \{ \Theta(u) : u \in \mathbb{R}_+^m \} \\ &= z_{LD}. \end{aligned}$$

Furthermore, $z_{SD} \leq z_{SP} \leq z$ holds directly from weak conic duality and Proposition 5. \square

From Theorem 2, the Lagrangian dual optimal value has no duality gap if and only if conic program itself has no duality gap and CPSD tensor relaxation is tight.

3.3 CP and CPSD tensor relaxations for POP with nonnegative variables

Consider following conic programs:

$$\begin{aligned} z_{CP} &= \inf \quad \langle T_d(p_0), X \rangle \\ \text{s.t.} \quad & \langle T_d(p_i), X \rangle \leq 0, \quad i = 1, \dots, m, \\ & \langle T_d(1), X \rangle = 1, \\ & X \in \mathcal{C}_{n+1,d}^*(\mathbb{R}_+^{n+1}), \end{aligned} \tag{10}$$

and

$$\begin{aligned} z_{SP,+} &= \inf \quad \langle T_d(p_0), X \rangle \\ \text{s.t.} \quad & \langle T_d(p_i), X \rangle \leq 0, \quad i = 1, \dots, m, \\ & \langle T_d(-x_i), X \rangle \leq 0, \quad i = 1, \dots, n, \\ & \langle T_d(1), X \rangle = 1, \\ & X \in \mathcal{C}_{n+1,d}^*(\mathbb{R}^{n+1}), \end{aligned} \tag{11}$$

and their conic dual problems

$$\begin{aligned}
 z_{CD} &= \sup \left\{ \mu : T_d(p_0) - \mu T_d(1) + \sum_{i=1}^m u_i T_d(p_i) \in \mathcal{C}_{n+1,d}(\mathbb{R}_+^{n+1}), u \in \mathbb{R}_+^m \right\}. \\
 z_{SD,+} &= \sup \left\{ \mu : T_d(p_0 - \mu) + \sum_{i=1}^m u_i T_d(p_i) \right. \\
 &\quad \left. + \sum_{i=1}^n v_i T_d(-x_i) \in \mathcal{C}_{n+1,d}(\mathbb{R}^{n+1}), u \in \mathbb{R}_+^m, v \in \mathbb{R}_+^n \right\}. \tag{12}
 \end{aligned}$$

Proposition 6 *Problem (10) and problem (11) are relaxations for problem (6) with $z_{CP} \leq z_+$ and $z_{SP,+} \leq z_+$.*

Theorem 3 *For problem (6), its Semi-Lagrangian dual function optimal value and its Lagrangian dual function optimal value satisfy*

$$\begin{aligned}
 z_{semi} &= \sup \left\{ \mu : (\mu, u) \in \mathbb{R} \times \mathbb{R}_+^m, T_d(L(x; u) - \mu) \in \mathcal{C}_{n+1,d}(\mathbb{R}_+^{n+1}) \right\}, \\
 z_{LD,+} &= \sup \left\{ \mu : (\mu, u, v) \in \mathbb{R} \times \mathbb{R}_+^m \times \mathbb{R}_+^n, T_d(L_+(x; u, v) - \mu) \in \mathcal{C}_{n+1,d}(\mathbb{R}^{n+1}) \right\},
 \end{aligned}$$

and

- (a) $z_{LD,+} \leq z_{semi} = z_{CD} \leq z_{CP} \leq z_+$.
- (b) $z_{LD,+} = z_{SD,+} \leq z_{SP,+} \leq z_+$.

Proof By Corollary 1,

$$\begin{aligned}
 \Theta_{semi}(u) &= \inf \{ L(x; u) : x \in \mathbb{R}^n \} \\
 &= \sup \left\{ \mu : T_d(L_+(x; u) - \mu) \in \mathcal{C}_{n+1,d}(\mathbb{R}^{n+1}) \right\}, \\
 \Theta_+(u, v) &= \inf \{ L_+(x; u, v) : x \in \mathbb{R}^n \} \\
 &= \sup \left\{ \mu : T_d(L_+(x; u, v) - \mu) \in \mathcal{C}_{n+1,d}(\mathbb{R}^{n+1}) \right\},
 \end{aligned}$$

then

$$\begin{aligned}
 z_{semi} &= \sup \{ \Theta_{semi}(u) : u \in \mathbb{R}_+^m \} \\
 &= \sup \left\{ \mu : (\mu, u) \in \mathbb{R} \times \mathbb{R}_+^m, T_d(L(x; u) - \mu) \in \mathcal{C}_{n+1,d}(\mathbb{R}_+^{n+1}) \right\}. \\
 z_{LD,+} &= \sup \{ \Theta_+(u, v) : u \in \mathbb{R}_+^m, v \in \mathbb{R}_+^n \} \\
 &= \sup \left\{ \mu : (\mu, u, v) \in \mathbb{R} \times \mathbb{R}_+^m \times \mathbb{R}_+^n, T_d(L_+(x; u, v) - \mu) \in \mathcal{C}_{n+1,d}(\mathbb{R}^{n+1}) \right\}.
 \end{aligned}$$

For (a), from (12), we have,

$$\begin{aligned}
 z_{CD} &= \sup \left\{ \mu : T_d(p_0) - \mu T_d(1) + \sum_{i=1}^m u_i T_d(p_i) \in \mathcal{C}_{n+1,d}(\mathbb{R}_+^{n+1}), u \in \mathbb{R}_+^m \right\} \\
 &= \sup \left\{ \mu : T_d(p_0) + \sum_{i=1}^m u_i p_i - \mu \in \mathcal{C}_{n+1,d}(\mathbb{R}_+^{n+1}), u \in \mathbb{R}_+^m \right\} \\
 &= \sup \left\{ \mu : (\mu, u) \in \mathbb{R} \times \mathbb{R}_+^m, T_d(L(x; u) - \mu) \in \mathcal{C}_{n+1,d}(\mathbb{R}_+^{n+1}) \right\} \\
 &= \sup \{ \Theta_{semi}(u) : u \in \mathbb{R}_+^m \} \\
 &= z_{semi}.
 \end{aligned}$$

And $z_{CD} \leq z_{CP} \leq z_+$ is an immediate result of weak conic duality and Proposition 6. For (b), from (12), we have

$$\begin{aligned} z_{SD,+} &= \sup \left\{ \mu : T_d(p_0 - \mu) + \sum_{i=1}^m u_i T_d(p_i) \right. \\ &\quad \left. + \sum_{i=1}^n v_i T_d(-x_i) \in C_{n+1,d}(\mathbb{R}^{n+1}), u \in \mathbb{R}_+^m, v \in \mathbb{R}_+^n \right\} \\ &= \sup \left\{ \mu : T_d \left(p_0(x) + \sum_{i=1}^m u_i p_i(x) \right. \right. \\ &\quad \left. \left. - \sum_{i=1}^n v^T x - \mu \right) \in C_{n+1,d}(\mathbb{R}^{n+1}), u \in \mathbb{R}_+^m, v \in \mathbb{R}_+^n \right\} \\ &= \sup \{ \mu : (\mu, u, v) \in \mathbb{R} \times \mathbb{R}_+^m \times \mathbb{R}_+^n, T_d(L_+(x; u, v) - \mu) \in C_{n+1,d}(\mathbb{R}^{n+1}) \} \\ &= \sup \{ \Theta_+(u, v) : u \in \mathbb{R}_+^m, v \in \mathbb{R}_+^n \} \\ &= z_{LD,+}. \end{aligned}$$

And $z_{SD,+} \leq z_{SP,+} \leq z_+$ holds directly from weak conic duality and Proposition 6. □

4 Quadratic reformulation for POPs and its relaxations

In Sect. 3, we show that CP and CPSD tensor relaxations are as tight as Lagrangian relaxations for general POPs. In this section, we will compare CP and CPSD tensor relaxations of a POP with CP and PSD matrix relaxations of the *quadratic reformulation* of the POP. By quadratic reformulation, we refer to the reformulation of the POP obtained by introducing additional variables and constraints so that all the polynomials involved in the POP can be rewritten as quadratic polynomials on the original and additional variables. Given a QCQP reformulation of a POP, an approximate solution to the POP can be obtained using well-studied SDP or CP relaxations. On the other hand, as discussed in Sect. 3, relaxations for general POPs can be obtained directly by using the CP or the CPSD tensor cones. In general, it is difficult to compare these two relaxation approaches. However, we show in this section that for a fairly wide class of POPs, the tensor relaxation approach provides bounds for the POP that are as tight as the bounds for the POP obtained by using the quadratic reformulation approach described above.

Note that the results stated thus far in the article hold in similar fashion for maximization problems (beyond minimization problems). In what follows in this section, we purposely choose to consider maximization POPs for ease of presentation of Theorem 4, the main result in this section.

4.1 QCQP reformulation of a POP

A general POP can be reformulated as a QCQP in different ways by adding appropriate additional variables and constraints (see, e.g., [43], Sect. 4.5) In this section, the main focus is on some classes of 4th degree POPs. Thus, we use a specific reformulation approach for such problems; that is, we will introduce additional variables to represent the quadratic terms (i.e. the square of single variable and the multiplication of two variables) of the variables in the original problem. Specifically, consider the following POP:

$$\begin{aligned}
 & \sup && p_0(x) \\
 & \text{s.t.} && p_i(x) \leq d_i, \quad i = 1, \dots, m_0, \\
 & && q_j(x) \leq 0, \quad j = 1, \dots, m_1, \\
 & && x \in \mathbb{R}_+^n,
 \end{aligned} \tag{13}$$

where $p_0(x) \in \mathbb{R}_4[x], q_j(x) \in \mathbb{R}_2[x]$ (recall that $\mathbb{R}_d[x] := \{p \in \mathbb{R}[x] : \deg(p) \leq d\}$) and $p_i(x)$ are homogeneous polynomials of degree 4. Problem (13) encompasses a large class of 4th degree optimization problems, including QCQPs. Some problems that belong to this class are biquadratic assignment problems [38,46], alternating current optimal power flow (ACOPF) problems [11,23,28,32], independent component analysis problems [16], blind channel equalization problems in digital communication [37] and sensor localization problems [7].

Define an index set

$$S = \left\{ (a, b, c) \in \mathbb{N}^3 : a = 1, \dots, n, b = a, \dots, n, c = \left(n - \frac{a}{2}\right)(a - 1) + b \right\}, \tag{14}$$

as the index for the additional variables to be added, where the index c goes from 1 to $|S| = \binom{n+1}{2}$, which is the maximum number of additional variables needed to reformulate 4th degree polynomials using 2nd degree polynomials. Specifically, introducing additional variables $y_c = x_a x_b, \forall (a, b, c) \in S$, the QCQP reformulation of problem (13) can be reformulated as

$$\begin{aligned}
 & \sup && q_0(x, y) \\
 & \text{s.t.} && h_i(y) \leq d_i, \quad i = 1, \dots, m_0, \\
 & && q_j(x) \leq 0, \quad j = 1, \dots, m_1, \\
 & && y_c - x_a x_b = 0, \quad \forall (a, b, c) \in S, \\
 & && x \in \mathbb{R}_+^n, y \in \mathbb{R}_+^{|S|},
 \end{aligned} \tag{15}$$

where $q_0(x, y)$ and $h_i(y), i = 1, \dots, m_0$, are the reformulated quadratic polynomials after replacing $x_a x_b$ with $y_c, \forall (a, b, c) \in S$ in $q_0(x), p_i(x), i = 1, \dots, m_0$. Clearly, problem (13) and (15) are equivalent. Furthermore, as $p_i(x)$ and $h_i(y)$ are homogeneous polynomials, then it follows that

$$\tilde{p}_i(x) = p_i(x) = h_i(y) = \tilde{h}_i(y), i = 1, \dots, m_0. \tag{16}$$

This fact will be used in Theorem 4 later in this section.

For ease of notation, let $z = (x, y) \in \mathbb{R}_+^{n+|S|}$, then (15) is equivalent to

$$\begin{aligned}
 & \sup && q_0(z) \\
 & \text{s.t.} && h_i(z) \leq d_i, \quad i = 1, \dots, m_0, \\
 & && q_j(z) \leq 0, \quad j = 1, \dots, m_1, \\
 & && z_{n+c} - z_a z_b = 0, \quad \forall (a, b, c) \in S, \\
 & && z \in \mathbb{R}_+^{n+|S|}.
 \end{aligned} \tag{17}$$

As an illustration of the quadratic reformulation discussed above, consider the next example.

Example 1 (QCQP reformulation) Consider the following univariate program,

$$\begin{aligned}
 \sup \quad & x^4 + x^3 + x^2 + x + 1 \\
 \text{s.t.} \quad & x^4 \leq 1, \\
 & x^2 - x - 2 \leq 0, \\
 & -x + 1 \leq 0, \\
 & x \in \mathbb{R}_+.
 \end{aligned} \tag{18}$$

Let $y = x^2$ and $z = (x, y) \in \mathbb{R}_+^2$, then problem (18) is equivalent to

$$\begin{array}{ll}
 \sup & y^2 + xy + y + x + 1 \\
 \text{s.t.} & y^2 \leq 1, \\
 & y - x - 2 \leq 0, \\
 & -x + 1 \leq 0, \\
 & y - x^2 = 0, \\
 & x \in \mathbb{R}_+, y \in \mathbb{R}_+.
 \end{array}
 \quad \equiv \quad
 \begin{array}{ll}
 \sup & z_2^2 + z_1z_2 + z_2 + z_1 + 1 \\
 \text{s.t.} & z_2 \leq 1, \\
 & z_2 - z_1 - 2 \leq 0, \\
 & -z_1 + 1 \leq 0, \\
 & z_2 - z_1^2 = 0, \\
 & z \in \mathbb{R}_+^2.
 \end{array}$$

4.2 CP matrix relaxations for a QCQP

Consider the following CP matrix relaxations for problem (17),

$$\begin{aligned}
 \sup \quad & \langle T_2(q_0(z)), Z \rangle \\
 \text{s.t.} \quad & \langle T_2(h_i(z)), Z \rangle \leq d_i, \quad i = 1, \dots, m_0, \\
 & \langle T_2(q_j(z)), Z \rangle \leq 0, \quad j = 1, \dots, m_1, \\
 & \langle T_2(1), Z \rangle = 1, \\
 & Z_{1,n+c+1} - Z_{a+1,b+1} = 0, \quad \forall (a, b, c) \in S, \\
 & Z \in \mathcal{C}_{n+|S|+1,2}^* \left(\mathbb{R}_+^{n+|S|+1} \right).
 \end{aligned} \tag{19}$$

After relaxing the equality constraints $Z_{1,c+n+1} - Z_{a+1,b+1} = 0, \forall (a, b, c) \in S$ into inequality constraints, we have the following CP tensor relaxation of problem (19)

$$\begin{aligned}
 \sup \quad & \langle T_2(q_0(z)), Z \rangle \\
 \text{s.t.} \quad & \langle T_2(h_i(z)), Z \rangle \leq d_i, \quad i = 1, \dots, m_0, \\
 & \langle T_2(q_j(z)), Z \rangle \leq 0, \quad j = 1, \dots, m_1, \\
 & \langle T_2(1), Z \rangle = 1, \\
 & Z_{1,c+n+1} - Z_{a+1,b+1} \leq 0, \quad \forall (a, b, c) \in S, \\
 & Z \in \mathcal{C}_{n+|S|+1,2}^* \left(\mathbb{R}_+^{n+|S|+1} \right),
 \end{aligned} \tag{20}$$

Proposition 7 *If problem (19) is feasible and the coefficients of $q_0(z)$ in problem (19) are nonnegative, then problems (19) and (20) are equivalent.*

Proof This follows from the fact that problem (20) is a relaxation of problem (19), and the fact that if the coefficients of $q_0(z)$ are nonnegative, then any optimal solution $Z \in \mathcal{C}_{n+|S|+1,2}^* \left(\mathbb{R}_+^{n+|S|+1} \right)$ of problem (20) would satisfy $Z_{1,c+n+1} = Z_{a+1,b+1}, \forall (a, b, c) \in S.$ □

Recall that the CP tensor relaxation (10) for general POPs. Below, we apply it directly to problem (17) to obtain the following conic program,

$$\begin{aligned}
 &\sup \quad \langle T_4(p_0(x)), X \rangle \\
 &\text{s.t.} \quad \langle T_4(p_i(x)), X \rangle \leq d_i, \quad i = 1, \dots, m_0, \\
 &\quad \quad \langle T_4(q_j(x)), X \rangle \leq 0, \quad j = 1, \dots, m_1, \\
 &\quad \quad \langle T_4(1), X \rangle = 1, \\
 &\quad \quad X \in C_{n+1,4}^* \left(\mathbb{R}_+^{n+1} \right),
 \end{aligned} \tag{21}$$

Problem (19) and (21) can be seen as two different relaxations for the POP (13). In problem (19) the polynomials with degree higher than 2 are reformulated as quadratic polynomials. SDP and CP matrix relaxations for the reformulated QCQP are well studied in the literature [3, 8–10, 13, 15, 27, 49, among others]. However, the introduction of the additional variables and constraints in problem (19) may result in the problem not satisfying the conditions required for QCQPs to be equivalent to their associated CP relaxation, even when the original POP satisfies the conditions required for its CP tensor relaxation being equivalent to the POP. Also, the additional variables and constraints can become substantially burdensome in terms of the problem’s size. In contrast, in problem (21) the polynomials with degree higher than 2 are represented by higher order tensors which avoids introducing additional variables and constraints. Next we will show that under some conditions, the latter relaxation will provide bounds that are at least as tight as the ones obtained with the former approach for problem (13).

Lemma 1 ([43, Lemma 2]) *For any $d > 0$ and $n > 0$, $C_{n+1,d}^*(\mathbb{R}_+^{n+1}) = \text{conic}(M_d(\{0, 1\} \times \mathbb{R}_+^n))$.*

Theorem 4 *Consider a feasible problem (13) where the coefficients of $p_0(x)$ are nonnegative, then problem (19) is a relaxation of problem (21).*

Proof By Proposition 7, problems (19) and (20) are equivalent. Using Lemma 1, for any feasible solution $X \in C_{n+1,4}^*(\mathbb{R}_+^{n+1})$ to problem (21) we have

$$X = \sum_{s=1}^{n_1} \lambda_s M_4(1, u_s) + \sum_{t=1}^{n_0} \gamma_t M_4(0, v_t),$$

for some $n_0, n_1 \geq 0$, $\lambda_s, \gamma_t > 0$ and $u_s, v_t \in \mathbb{R}_+^n$. Then by using (1),

$$\begin{aligned}
 1 &= \langle T_4(1), X \rangle = \sum_{s=1}^{n_1} \lambda_s, \\
 d_i &\geq \langle T_4(p_i), X \rangle = \sum_{s=1}^{n_1} \lambda_s p_i(u_s) + \sum_{t=1}^{n_0} \gamma_t \tilde{p}_i(v_t), \quad i = 1, \dots, m_0, \\
 0 &\geq \langle T_4(q_j), X \rangle = \sum_{s=1}^{n_1} \lambda_s q_j(u_s) + \sum_{t=1}^{n_0} \gamma_t \tilde{q}_j(v_t), \quad j = 1, \dots, m_1,
 \end{aligned} \tag{22}$$

with an objective function value of $\sum_{s=1}^{n_1} \lambda_s p_0(u_s) + \sum_{t=1}^{n_0} \gamma_t \tilde{p}_0(v_t)$. Recall the index set S in (14), and construct a vector of w_s, w'_t for $s = 1, \dots, n_1, t = 1, \dots, n_0$ as follows:

$$\begin{aligned}
 (w_s)_c &= (u_s)_a (u_s)_b, \quad (a, b, c) \in S, \\
 (w'_t)_c &= (v_t)_a (v_t)_b, \quad (a, b, c) \in S.
 \end{aligned} \tag{23}$$

Next we show that

$$Z = \sum_{s=1}^{n_1} \lambda_s M_2(1, (u_s, w_s)) + \sum_{t=1}^{n_0} \gamma_t M_2(0, (v_t, w'_t)), \tag{24}$$

is a feasible solution to problem (19). Clearly, $Z \in C_{n+|S|+1,2}^*(\mathbb{R}_+^{n+|S|+1})$, and from equation (23) and (24), we have

$$\begin{aligned} Z_{1,c+n+1} &= \sum_{s=1}^{n_1} \lambda_s (w_s)_c = \sum_{s=1}^{n_1} \lambda_s (u_s)_a (u_s)_b, \forall (a, b, c) \in S, \\ Z_{a+1,b+1} &= \sum_{s=1}^{n_1} \lambda_s (u_s)_a (u_s)_b + \sum_{t=1}^{n_0} \gamma_t (v_t)_a (v_t)_b, \forall (a, b, c) \in S, \end{aligned}$$

which indicates that $Z_{1,c+n+1} \leq Z_{a+1,b+1}, \forall (a, b, c) \in S$. From (22), it follows that

$$\begin{aligned} \langle T_2(1), Z \rangle &= \sum_{s=1}^{n_1} \lambda_s = 1, \\ \langle T_2(q_j), Z \rangle &= \sum_{s=1}^{n_1} \lambda_s q_j(u_s) + \sum_{t=1}^{n_0} \gamma_t \tilde{q}_j(v_t) \leq 0, \quad j = 1, \dots, m_1. \end{aligned}$$

Also, given that $p_i(x), i = 1, \dots, m_0$ are homogeneous polynomials of degree 4, it follows from equations (16) and (22) that

$$\begin{aligned} \langle T_2(h_i), Z \rangle &= \sum_{s=1}^{n_1} \lambda_s h_i(w_s) + \sum_{t=1}^{n_0} \gamma_t \tilde{h}_i(w'_t) \\ &= \sum_{s=1}^{n_1} \lambda_s p_i(u_s) + \sum_{t=1}^{n_0} \gamma_t \tilde{p}_i(v_t) \leq d_i, \quad i = 1, \dots, m_0. \end{aligned} \tag{25}$$

Furthermore, the feasible solution Z has an objective value equal to

$$\sum_{s=1}^{n_1} \lambda_s q_0(u_s, w_s) + \sum_{t=1}^{n_0} \gamma_t \tilde{q}_0(v_t, w'_t) = \sum_{s=1}^{n_1} \lambda_s p_0(u_s) + \sum_{t=1}^{n_0} \gamma_t \tilde{q}_0(v_t, w'_t).$$

Under the condition that $p_0(x)$ has nonnegative coefficients and $x \in \mathbb{R}_+^n$, we have

$$\sum_{t=1}^{n_0} \gamma_t \tilde{q}_0(v_t, w'_t) \geq \sum_{t=1}^{n_0} \gamma_t \tilde{p}_0(v_t).$$

Therefore, from any feasible solution to problem (21), we can construct a feasible solution to problem (20) with a larger objective function value, which indicates that problem (19) is a relaxation for problem (21). \square

To illustrate the use of the condition that $p_i(x), i = 1, \dots, m_0$ are homogeneous polynomials of degree 4 in Theorem 4, consider a constraint with two variables, $p_1(x_1, x_2) = x_1^4 + x_1^2 x_2 \leq 1, x_1, x_2 \geq 0$. For the QCQP reformulation of the constraint, the additional variable $y = x_1^2$ is introduced. Then $h_1(x_1, x_2, y) = y^2 + yx_2$. Thus

$$\tilde{h}_1(x_1, x_2, y) = y^2 + yx_2 \geq y^2 = x_1^4 = \tilde{p}_1(x_1, x_2),$$

and it indicates (16) might not hold, which is needed for the proof of Theorem 4. Also note that for minimization POPs, Theorem 4 applies after changing the condition on the polynomial $p_0(x)$ to have nonpositive coefficients.

Theorem 4 proves that the CP tensor relaxation can provide bounds that are at least as tight as the ones obtained from CP matrix relaxations of a quadratic reformulation of a class of POPs. In next section, we will provide the results of numerical experiments that show that indeed, bounds for a POP based on CP tensor relaxations are tighter than the ones obtained via CP matrix relaxations.

5 Numerical comparison of two relaxations for PO

Unlike the tractability of the PSD matrix cone, the CPSD tensor cone is not tractable in general. Also similar to the intractability of CP matrices, the CP tensor cone is also not tractable in general. In this section, we will discuss and develop tractable approximations for the CP and the CPSD tensor cones. Then, we use these approximations to show that these CP and CPSD tensor relaxation of a POP provide tighter bounds than CP and PSD relaxations of the QCQP obtained from a quadratic reformulation of the POP.

5.1 Approximation of the CP and CPSD tensor cones

Let us first introduce some additional notation. For $T = M_d(x)$, $x \in \mathbb{R}^n$, denote $T_{(i_1, \dots, i_d)}$ as the element in (i_1, \dots, i_d) position of tensor T , where $(i_1, \dots, i_d) \in \{1, \dots, n\}^d$. To be more specific, i_j with $j = 1, \dots, d$ means the choice of $\{x_1, \dots, x_n\}$ in the j th position in the tensor product, i.e. $i_1 = 2$ means choosing x_2 as the first position in the tensor product. To illustrate, let $x \in \mathbb{R}^3$ and let

$$T^1 = M_2(x) = \begin{pmatrix} x_1^2 & x_1x_2 & x_1x_3 \\ x_1x_2 & x_2^2 & x_2x_3 \\ x_1x_3 & x_2x_3 & x_3^2 \end{pmatrix},$$

then $T^1_{(1,2)} = x_1x_2$ and it is in the (1,2) position in T^1 . Also for $T = M_d(x)$, $x \in \mathbb{R}^n$, when $d > 2$, let $T_{(i_1, \dots, i_{d-2}, \cdot, \cdot)}$ denote the matrix whose elements are given by

$$(T_{(i_1, \dots, i_{d-2}, \cdot, \cdot)})_{jk} = T_{(i_1, \dots, i_{d-2}, j, k)}, \quad j, k = 1, \dots, n.$$

For example, let $T^2 = M_3(x)$, $x \in \mathbb{R}^3$, then

$$T^2_{(1, \cdot, \cdot)} = \begin{pmatrix} x_1^3 & x_1^2x_2 & x_1^2x_3 \\ x_1^2x_2 & x_1x_2^2 & x_1x_2x_3 \\ x_1^2x_3 & x_1x_2x_3 & x_1x_3^2 \end{pmatrix}, \quad T^2_{(2, \cdot, \cdot)} = \begin{pmatrix} x_1^2x_2 & x_1x_2^2 & x_1x_2x_3 \\ x_1x_2^2 & x_2^3 & x_2^2x_3 \\ x_1x_2x_3 & x_2^2x_3 & x_2x_3^2 \end{pmatrix}.$$

Definition 5 Let $T = M_d(x)$, $x \in \mathbb{R}^n$. For any $(i_1, \dots, i_{d-2}) \in \{1, \dots, n\}^{d-2}$, $T_{(i_1, \dots, i_{d-2}, \cdot, \cdot)}$ is a *principal matrix* if $I_k \subseteq \{0, \dots, d-2\}$ is even for all $k = 1, \dots, n$, where I_k is an ordered set of the number of appearance $i_j = k, \forall j = 1, \dots, d-2$ where $k = 1, \dots, n$.

For example, let $T^3 = M_8(x)$, $x \in \mathbb{R}^3$, then

$$T^3_{(1,1,2,2,3,3, \cdot, \cdot)}, T^3_{(1,2,2,2,1,2, \cdot, \cdot)}, T^3_{(2,3,2,1,3,1, \cdot, \cdot)} \text{ are principal matrices;}$$

$$T^3_{(1,1,1,2,3,3, \cdot, \cdot)}, T^3_{(1,2,2,2,2,2, \cdot, \cdot)}, T^3_{(2,3,2,2,3,1, \cdot, \cdot)} \text{ are not principal matrices.}$$

Next we will discuss the approximation strategies for the CP and the CPSD tensor cones based on PSD and DNN matrices.

Definition 6 A symmetric matrix X is called *doubly nonnegative (DNN)* if and only if $X \succeq 0$ and $X \geq 0$, where $X \geq 0$ indicates that every element of X is nonnegative.

Proposition 8 Let $T \in \mathcal{S}_{n,d}$ be a given symmetric tensor.

- (a) If $T \in \mathcal{C}_{n,d}^*(\mathbb{R}_+^n)$, then $T_{(i_1, \dots, i_d)} \geq 0, T_{(i_1, \dots, i_{d-2}, \cdot, \cdot)} \geq 0, \forall (i_1, \dots, i_d) \in \{1, \dots, n\}^d$.
- (b) If $T \in \mathcal{C}_{n,d}^*(\mathbb{R}^n)$, for all principal matrices $T_{(i_1, \dots, i_{d-2}, \cdot, \cdot)}, T_{(i_1, \dots, i_{d-2}, \cdot, \cdot)} \geq 0, \forall (i_1, \dots, i_d) \in \{1, \dots, n\}^d$.

Proof For part (a), by Proposition 3 (a), $T = \sum_{i \in \mathbb{N}} \lambda_i M_d(x^i)$, where $x^i \in \mathbb{R}_+^n, \lambda_i \geq 0$, for all $i \in \mathbb{N}$, and $\sum_{i \in \mathbb{N}} \lambda_i = 1$. Then it is clear that $T_{(i_1, \dots, i_d)} \geq 0$, and

$$T_{(i_1, \dots, i_{d-2}, \cdot, \cdot)} = \sum_{i \in \mathbb{N}} \lambda_i \prod_{k=1}^n (x_k^i)^{I_k} (x^i (x^i)^\top), \tag{26}$$

as $x^i (x^i)^\top \geq 0, \forall i \in \mathbb{N}$ and $\prod_{k=1}^n (x_k^i)^{I_k} \geq 0$, then $T_{(i_1, \dots, i_{d-2}, \cdot, \cdot)} \geq 0$, and $T_{(i_1, \dots, i_{d-2}, \cdot, \cdot)} \geq 0$ for all $(i_1, \dots, i_d) \in \{1, \dots, n\}^d$. For part (b), notice that the number of appearance $I_k, k = 1, \dots, n$ is even if $T_{(i_1, \dots, i_{d-2}, \cdot, \cdot)}$ is a principal matrix, then the proof follows as in part (a) from the fact that $\prod_{k=1}^n (x_k^i)^{I_k} \geq 0$ in (26). □

Example 2 To illustrate Proposition 8, take $T \in \mathcal{C}_{2,4}^*(\mathbb{R}_+^2)$ as an example. By Proposition 3

(a), $T = \sum_{i \in \mathbb{N}} \lambda_i M_4(x^i)$, where $\lambda_i \geq 0, \sum_i \lambda_i = 1$ and $x^i \in \mathbb{R}_+^2$. Then for any $y \in \mathbb{R}^2$,

$$y^\top T_{(1,2, \cdot, \cdot)} y = y^\top \sum_{i \in \mathbb{N}} \lambda_i M_4(x^i)_{(1,2, \cdot, \cdot)} y = x_1^i x_2^i \sum_{i \in \mathbb{N}} (y^\top x^i)^2 \geq 0,$$

which indicates that $T_{(1,2, \cdot, \cdot)}$ is a 2×2 positive semidefinite matrix.

Next we discuss linear matrix inequality (LMI) approximation of the CPSD and the CP tensor cones. Based on Proposition 8, define the following tensor cones

$$\begin{aligned} \mathcal{K}_{n,d}^{SDP} &= \left\{ T \in \mathcal{S}_{n,d} : T_{(i_1, \dots, i_{d-2}, \cdot, \cdot)} \geq 0, \forall (i_1, \dots, i_{d-2}) \in \{1, \dots, n\}^{d-2} \right\}, \\ \mathcal{K}_{n,d}^L &= \left\{ T \in \mathcal{S}_{n,d} : T_{(i_1, \dots, i_d)} \geq 0, \forall (i_1, \dots, i_d) \in \{1, \dots, n\}^d \right\}, \\ \mathcal{K}_{n,d}^{DNN} &= \left\{ T \in \mathcal{S}_{n,d} : T_{(i_1, \dots, i_{d-2}, \cdot, \cdot)} \geq 0, T_{(i_1, \dots, i_{d-2}, \cdot, \cdot)} \geq 0, \forall (i_1, \dots, i_{d-2}) \in \{1, \dots, n\}^{d-2} \right\}. \end{aligned} \tag{27}$$

It is easy to see these cones are *convex closed cones* satisfying

$$\begin{aligned} \mathcal{C}_{n,d}^*(\mathbb{R}^n) &\subseteq \mathcal{K}_{n,d}^{SDP}, \\ \mathcal{C}_{n,d}^*(\mathbb{R}_+^n) &\subseteq \mathcal{K}_{n,d}^{DNN} \subseteq \mathcal{K}_{n,d}^L. \end{aligned} \tag{28}$$

Consider the following conic program,

$$\begin{aligned} [\mathbf{TP} - \mathcal{K}] \quad &\inf \langle T_d(p_0), X \rangle \\ &\text{s.t. } \langle T_d(p_i), X \rangle \leq 0, \quad i = 1, \dots, m, \\ &\langle T_d(1), X \rangle = 1, \\ &X \in \mathcal{K}_{n+1,d}. \end{aligned}$$

Table 1 Size comparison of different relaxations for 4th-degree POPs

	PSD matrix size	PSD matrix number	Total number of variables
[QP _{SDP}]	$(1 + n + S) \times (1 + n + S)$	1	$O(n^4)$
[TP – \mathcal{K}^{SDP}]	$(1 + S) \times (1 + n)$	n	$O(n^3)$
[QP _{DNN}]	$(1 + n + S) \times (1 + n + S)$	1	$O(n^4)$
[TP – \mathcal{K}^{DNN}]	$(1 + n) \times (1 + n)$	$O(n^2)$	$O(n^4)$

From (28), problem [TP- \mathcal{K}] is a LMI relaxation for problem (8) when \mathcal{K} is one of the cones defined in (27). These relaxations allow to approximately solve general POPs. In particular, it follows that

$$z_{[\text{TP}-\mathcal{K}^{SDP}]} \leq z_{SP} \leq z,$$

$$z_{[\text{TP}-\mathcal{K}^L]} \leq z_{[\text{TP}-\mathcal{K}^{DNN}]} \leq z_{CP} \leq z_+.$$

5.2 Numerical results

In Sect. 5.1, several LMI approximations for the CP and the CPSD tensor cones have been proposed to provide tractable relaxations for CP and CPSD tensor programs. In this section, we will provide numerical results on more general POP cases (compared to Sect. 4.2) in order to compare the bounds of two relaxation approaches discussed in Sect. 4.2. Similar to [TP – \mathcal{K}^L] and [TP – \mathcal{K}^{DNN}], denote [QP_L] and [QP_{DNN}] as the linear relaxation and DNN relaxation for problem (19). Also, denote by [QP_{SDP}] the SDP relaxation for the quadratic reformulation of problem (7). In Table 1, we compare the two approaches in terms of number and size of PSD matrices for 4th-degree POPs, to show that both approaches result in the solution of similar sized SDPs.

Next we present some preliminary results on small scale POPs to illustrate the performance of CP and CPSD tensor relaxations. Note that only bounds are compared as the time differences are negligible for the small scale examples considered below. All the numerical experiments are conducted on a 2.4 GHz CPU laptop with 8 GB memory. We implement all the models with YALMIP [34] in Matlab. We use SeDuMi as the SDP solver and CPLEX as the LP solver. For examples in Sects. 5.2.4 and 5.2.5, we use Couenne as the global solver.

5.2.1 Simple POP

Consider the following problem,

$$\min \left(\sum_{i=1}^n x_i \right)^4 \tag{29}$$

s.t. $x_1^4 = 1,$
 $x_i \geq 0, i = 1, \dots, n.$

By observation, the optimal value is 1, with an optimal solution $x_1^* = 1, x_k^* = 0, k = 2, \dots, n.$ The QCQP reformulation of (29) with the least number of additional variables is

$$\begin{aligned}
 \min \quad & y_1^2 \\
 \text{s.t.} \quad & y_1 = \left(\sum_{i=1}^n x_i \right)^2, \\
 & y_2 = x_1^2, \\
 & y_2^2 = 1, \\
 & x_i \geq 0, i = 1, \dots, n, \\
 & y_1, y_2 \geq 0.
 \end{aligned} \tag{30}$$

Relaxation [QP_L] for (30) gives an optimal value of 0. Relaxation [TP – K^L] can be directly applied to (29) and gives an optimal value of 1, which means the approximation by using tensor relaxation is tight.

5.2.2 Bi-quadratic POPs

Bi-quadratic problem and its difficulty have been studied in [33]. Consider the following specific bi-quadratic POP,

$$\begin{aligned}
 \min_{x \in \mathbb{R}^n, y \in \mathbb{R}^m} \quad & p_0 := \sum_{\substack{1 \leq i < j \leq n \\ 1 \leq a < b \leq m}} x_i x_j y_a y_b \\
 \text{s.t.} \quad & \|x\|^2 = 1, \|y\|^2 = 1,
 \end{aligned} \tag{31}$$

where $\|\cdot\|$ is the standard Euclidean norm. It is clear that problem (31) is equivalent to

$$\begin{aligned}
 \min_{x \in \mathbb{R}^n, y \in \mathbb{R}^m} \quad & \frac{1}{4} [x^\top (e_n e_n^\top - I_n) x] [y^\top (e_m e_m^\top - I_m) y] \\
 \text{s.t.} \quad & \|x\|^2 = 1, \|y\|^2 = 1,
 \end{aligned}$$

where e_n, e_m are the all-ones vectors of dimension n and m respectively, and I_n, I_m are the identity matrices of dimension $n \times n$ and $m \times m$. It is clear that the optimal value is $-\frac{1}{4}(\max\{n, m\} - 1)$. By defining an index set

$$\begin{aligned}
 S(n) &= \{(i, j, k) \in \mathbb{N}^3 : i = 1, \dots, n - 1, j = i + 1, \dots, n, k \\
 &= \left(n - \frac{i}{2} \right) (i - 1) + j - i \},
 \end{aligned}$$

we can reformulate problem (31) as a QCQP by introducing appropriate additional variables and constraints as

$$\begin{aligned}
 \min \quad & \sum_{\substack{1 \leq k \leq |S(n)| \\ 1 \leq c \leq |S(m)|}} w_k z_c \\
 \text{s.t.} \quad & w_k = x_i x_j, \forall (i, j, k) \in S(n), \\
 & z_c = y_a y_b, \forall (a, b, c) \in S(m), \\
 & \|x\|^2 = 1, \|y\|^2 = 1,
 \end{aligned} \tag{32}$$

where $w, z \in \mathbb{R}^m$ with $|S(n)| = n(n - 1)/2, |S(m)| = m(m - 1)/2$. Let $u = (x, y, w, z)$, then a naive SPD relaxation of problem (32) is given by

$$\begin{aligned}
 & \min \quad \sum_{\substack{n+m+1 \leq p \leq n+m+|S(n)| \\ n+m+|S(n)|+1 \leq q \leq n+m+|S(n)|+|S(m)|}} Q_{pq} \\
 & \text{s.t.} \quad u_{n+m+k} = Q_{ij}, \forall (i, j, k) \in S(n), \\
 & \quad u_{n+m+|S(n)|+c} = Q_{n+a,n+b}, \forall (a, b, c) \in S(m), \\
 & \quad \sum_{i=1}^n Q_{ii} = 1, \\
 & \quad \sum_{i=n+1}^{n+m} Q_{ii} = 1, \\
 & \quad \begin{pmatrix} 1 & u^T \\ u & Q \end{pmatrix} \in \mathcal{C}_{n+m+|S(n)|+|S(m)|+1,2}^* \left(\mathbb{R}^{n+m+|S(n)|+|S(m)|+1} \right).
 \end{aligned} \tag{33}$$

Although we use (33) in the analysis that follows, it is worth to mention that more elaborated SDP relaxations of (31), that provide bounds with guaranteed performance, are discussed in [33].

Proposition 9 *Problem (33) is unbounded.*

Proof Let \bar{u} be a $(n + m + |S(n)| + |S(m)|) \times 1$ all-zero vector and let \bar{Q} be a $(n + m + |S(n)| + |S(m)|) \times (n + m + |S(n)| + |S(m)|)$ matrix such that

$$\begin{aligned}
 \bar{Q}_{11} = \bar{Q}_{n+1,n+1} = 1, \quad \bar{Q}_{n+m+1,n+m+1} = \bar{Q}_{n+m+|S(n)|+1,n+m+|S(n)|+1} = M^2, \\
 \bar{Q}_{n+m+1,n+m+|S(n)|+1} = \bar{Q}_{n+m+|S(n)|+1,n+m+1} = -M,
 \end{aligned}$$

where M is a positive number and let all other entries for \bar{Q} be 0. It is clear that (\bar{u}, \bar{Q}) is a feasible solution to problem (33). However, as $M \rightarrow \infty$, the objective function values goes to $-\infty$, thus the problem is unbounded. □

Proposition 9 shows that the relaxation $[\mathbf{QP}_{SDP}]$ for problem (31) fails to provide a bound. However, a CPSD tensor cone can be directly applied to problem (31),

$$\begin{aligned}
 & \min \quad \langle T_4(p_0), X \rangle \\
 & \text{s.t.} \quad \langle T_4(\|x\|^2), X \rangle = 1, \\
 & \quad \langle T_4(\|y\|^2), X \rangle = 1, \\
 & \quad \langle T_4(1), X \rangle = 1, \\
 & \quad X \in \mathcal{C}_{n+m+1,4}^* \left(\mathbb{R}^{n+m+1} \right),
 \end{aligned} \tag{34}$$

where $p_0(x)$ is the objective function of problem (31).

In Table 2, we can see that the relaxation $[\mathbf{TP} - \mathcal{K}^{SDP}]$ can provide the optimal value for problem (33) while relaxation $[\mathbf{QP}_{SDP}]$ for the QCQP reformulation of problem (33) fails to give a bound.

Table 2 Relaxation Comparisons for Example 5.2.2

(n, m)	Optimal	$[\mathbf{TP} - \mathcal{K}^{SDP}]$	(n, m)	Optimal	$[\mathbf{TP} - \mathcal{K}^{SDP}]$
(2,2)	-0.25	-0.25	(2, 10)	-2.25	-2.25
(3,3)	-0.50	-0.50	(3, 9)	-2.00	-2.00
(4,4)	-0.75	-0.75	(4, 8)	-1.75	-1.75
(5,5)	-1.00	-1.00	(5, 7)	-1.50	-1.50
(6,6)	-1.25	-1.25			
(7,7)	-1.50	-1.50			
(8,8)	-1.75	-1.75			
(9,9)	-2.00	-2.00			
(10,10)	-2.25	-2.25			

5.2.3 Non-convex QCQP

Consider the following nonconvex QCQP,

$$\begin{aligned}
 \min \quad & f_0(x) = -8x_1^2 - x_1x_2 - 13x_2^2 - 6x_1 - x_2 \\
 \text{s.t.} \quad & f_1(x) = x_1^2 + x_1x_2 + 2x_2^2 - 3x_1 - 3x_2 - 7 \leq 0, \\
 & f_2(x) = 2x_1x_2 + 33x_1 + 15x_2 - 10 \leq 0, \\
 & f_3(x) = x_1 + 2x_2 - 6 \leq 0, \\
 & x_1, x_2 \geq 0.
 \end{aligned} \tag{35}$$

The optimal solution of problem (35) is $x^* = (0, 0.6667)^\top$ with $f_0(x^*) = -6.4444$ (see [51]). A PSD relaxation and a CP relaxation of (35) have been studied in [51], that give bounds of -103.43 and -26.67 respectively for problem (35) (refer to Table 2 in [51], where (SDP+RLT) is actually a *DNN* relaxation for the CP relaxation of (35)).

Now consider the equivalent formulation of (35), obtained after adding the valid inequalities $x_2 f_2(x) \leq 0, x_1^2 f_1(x) \leq 0$:

$$\begin{aligned}
 \min \quad & f_0(x) = -8x_1^2 - x_1x_2 - 13x_2^2 - 6x_1 - x_2 \\
 \text{s.t.} \quad & f_1(x) = x_1^2 + x_1x_2 + 2x_2^2 - 3x_1 - 3x_2 - 7 \leq 0, \\
 & f_2(x) = 2x_1x_2 + 33x_1 + 15x_2 - 10 \leq 0, \\
 & f_3(x) = x_1 + 2x_2 - 6 \leq 0, \\
 & x_2 f_2(x) \leq 0, \\
 & x_1^2 f_1(x) \leq 0, \\
 & x_1, x_2 \geq 0.
 \end{aligned} \tag{36}$$

The QCQP reformulation of (36) obtained by adding appropriate additional variables and constraints is given by

Table 3 Relaxation comparisons for problem (35)

	Without Valid Inequalities		With Valid Inequalities		With PSD–RLT
	[QP _{SDP}]	[QP _{DNN}]	[QP _{DNN}]	[TP – \mathcal{K}^{DNN}]	[TP – \mathcal{K}^{DNN}]
Bound	-103.43	-26.67	-26.67	-12.83	-6.44 ^a

^a Optimal value is obtained

$$\begin{aligned}
 \min \quad & f_0(x) = -8x_1^2 - x_1x_2 - 13x_2^2 - 6x_1 - x_2 \\
 \text{s.t.} \quad & -y_1 = x_1^2 + x_1x_2 + 2x_2^2 - 3x_1 - 3x_2 - 7 \leq 0, \\
 & -y_2 = 2x_1x_2 + 33x_1 + 15x_2 - 10 \leq 0, \\
 & f_3(x) = x_1 + 2x_2 - 6 \leq 0, \\
 & y_3 = x_1^2, \\
 & -x_2y_2 \leq 0, \\
 & -y_1y_3 \leq 0, \\
 & x_1, x_2, y_1, y_2, y_3 \geq 0.
 \end{aligned}$$

After using the [QP_{DNN}] relaxation on problem (36) with valid inequalities, the original bound of -26.67 obtained from the [QP_{DNN}] relaxation on problem (35) is not improved. In contrast, the [TP – \mathcal{K}^{DNN}] tensor relaxation on problem (36) provides a tighter bound, -12.83, for the optimal value of (35).

In addition to adding valid inequalities discussed above, adding valid PSD constraints based on the *reformulation linearization technique* (RLT) can further strengthen the relaxations. Similar to the second order RLT-based valid constraints introduced in [14], using the constraint $\langle T_4(1), X \rangle = 1$, the conic constraint $X \in \mathcal{C}_{3,4}^*(\mathbb{R}^3)$, and a quadratic constraint $c_0 + c_{10}x_1 + c_{01}x_2 + c_{11}x_1^2 + c_{12}x_1x_2 + c_{22}x_2^2 \geq 0$, the following valid PSD–RLT constraints for the CP tensor relaxation of problem (36) can be constructed

$$\begin{aligned}
 & (c_0 + c_{10}x_1 + c_{01}x_2 + c_{11}x_1^2 + c_{12}x_1x_2 + c_{22}x_2^2) X_{(0,0,\dots)} \\
 & = c_0X_{(0,0,\dots)} + c_{10}X_{(1,0,\dots)} + c_{01}X_{(2,0,\dots)} + c_{11}X_{(1,1,\dots)} \\
 & + c_{12}X_{(1,2,\dots)} + c_{22}X_{(2,2,\dots)} \geq 0,
 \end{aligned} \tag{37}$$

where $X_{(0,0,\dots)} \geq 0$ as discussed in Sect. 5.1. Note that the CP tensor relaxations allow for the straightforward use of the valid PSD–RLT constraints. More importantly, with the valid PSD–RLT constraints, the [TP – \mathcal{K}^{DNN}] tensor relaxation gives the optimal value -6.4444 of problem (35).

A summary of the numerical results on problem (35) is given in Table 3.

5.2.4 Random objective function with ellipsoidal feasible region

Here, we present numerical results on randomly generated 4th degree POPs with an ellipsoidal feasible region. The test problem is

$$\begin{aligned}
 \min \quad & p_0(x_1, x_2, x_3) \\
 \text{s.t.} \quad & (x_1 - 0.5)^2 + (x_2 - 0.5)^2 + (x_3 - 0.5)^2 \geq 0.2^2, \\
 & (x_1 - 0.5)^2 + (x_2 - 0.5)^2 + (x_3 - 0.5)^2 \leq 0.6^2, \\
 & 0 \leq x_1, x_2, x_3 \leq 1,
 \end{aligned} \tag{38}$$

Table 4 Bound comparisons for problem (38)

Test No.	[QP _{DNN}]	[TP - K ^{DNN}]	ratio	[TP - K ^{DNN}] ⁺	Couenne
1	-4.6732	-3.2860	36.82%	-0.9055 ^a	-0.9055 ^a
2	-8.8748	-5.2725	78.15%	-4.2654 ^a	-4.2654 ^a
3	-5.6429	-4.1135	76.76%	-3.6477 ^a	-3.6477 ^a
4	-3.5507	-2.1173	53.59%	-0.8761 ^a	-0.8761 ^a
5	-11.0434	-9.5248	37.81%	-7.0268 ^a	-7.0268 ^a
6	-12.6822	-10.5600	24.46%	-4.0055 ^a	-4.0055 ^a
7	-3.0709	-2.4427	45.84%	-1.7005 ^a	-1.7005 ^a
8	0	0.0122	100%	0.0122 ^a	0.0122 ^a
9	-1	0.0091	100%	0.0091 ^a	0.0091 ^a
10	-5.2621	-1.9963	83.15%	-1.3345 ^a	-1.3345 ^a
11	-0.8450	-0.8438	0.30%	-0.4922 ^a	-0.4922 ^a
12	-3.5894	-2.9945	100%	-1.4597 ^a	-1.4597 ^a
13	-0.8554	-0.7762	11.70%	-0.1787 ^a	-0.1787 ^a
14	-6.1631	-2.6502	81.87%	-1.8723 ^a	-1.8723 ^a
15	-0.2666	-0.2666	0	-0.1487 ^a	-0.1487 ^a
16	-6.0238	-5.6216	10.16%	-2.0645 ^a	-2.0645 ^a
17	-4.9579	-4.9579	0	-4.0253 ^a	-4.0253 ^a
18	0	0.0080	100%	0.0080 ^a	0.0080 ^a
19	-12.1584	-10.7368	16.94%	-3.7659 ^a	-3.7659 ^a
20	-0.6545	0.0112	100%	0.0112 ^a	0.0112 ^a

^a Optimal value is obtained.
 [TP - K^{DNN}]⁺ := [TP - K^{DNN}] with PSD–RLT constraints

where $p_0(x_1, x_2, x_3)$ is a 4th degree polynomial whose coefficients are randomly selected from integers in the range $[-5, 5]$. The first constraint make the feasible region non-convex. Also, it is easy to see that the problem is feasible and bounded. We use the [TP - K^{DNN}] relaxation to approximate problem (38) and the [QP_{DNN}] relaxation to approximate the QCQP reformulation of problem (38). To compare these relaxation, and following [51], we use the following improvement ratio

$$ratio = \frac{[TP - K^{DNN}] - [QP_{DNN}]}{f_{opt} - [QP_{DNN}]},$$

where f_{opt} denotes the optimal objective value of problem (38).

We also add PSD–RLT constraints to problem (38) using the constraints $(x_1 - 0.5)^2 + (x_2 - 0.5)^2 + (x_3 - 0.5)^2 \geq 0.2^2$ and $(x_1 - 0.5)^2 + (x_2 - 0.5)^2 + (x_3 - 0.5)^2 \leq 0.6^2$. In Table 4, the relaxation [TP - K^{DNN}] with PSD–RLT constraints provides the tightest bounds, and in fact, the optimal value of the problems. The relaxation [TP - K^{DNN}] provides tighter bounds than [QP_{DNN}] for most test instances. For instances 8, 9, 18 and 20, the relaxation [TP - K^{DNN}] gives the optimal objective value, while [QP_{DNN}] is not tight. For instances 15 and 17, [TP - K^{DNN}] and [QP_{DNN}] give the same bound. An average of 50% improve ratio implies that [TP - K^{DNN}] has better performance [QP_{DNN}] in approximating problem (38).

Table 5 Relaxation Comparisons for Randomly Generated POPs

Test No.	Couenne	[$\mathbf{TP} - \mathcal{K}^{DNN}$]	[\mathbf{QP}_{DNN}]
1	-0.1790	-0.1852	Unbounded
2	10.9275	7.8888	0
3	-158.751	-245.7888	Unbounded
4	1.3041	1.1044	0
5	2.5418	1.9276	0
6	0.7107	-2.0031	Unbounded
7	1.0663	-6.6609	Unbounded
8	-8.0284	-56.0924	Unbounded
9	0.0275	0.0272	0
10	8.0032	2.4765	Unbounded

5.2.5 Numerical results on randomly generated POPs

Next, we present numerical results on randomly generated POPs. The objective function is a 4th degree homogenous polynomial on 3 variables, with two 4th degree polynomial inequality constraints, a linear inequality constraint and nonnegative variables. The coefficients in the objective function are integers in the range $[-5, 5]$ and the coefficients of the two polynomial constraints are integers in the range $[-10, 10]$ and the coefficients of linear constraint are integers in the range $[0, 5]$, with a right hand side coefficient in the range $[5, 15]$. We generate the problems and solve them with Couenne. For those problems which are feasible in Couenne, we use [$\mathbf{TP} - \mathcal{K}^{DNN}$] to directly approximate these problems and [\mathbf{QP}_{DNN}] to approximate the QCQP reformulation of these problems. Note that the convexity of these problems is not tested. Results are shown in Table 5, where we can clearly see that relaxation [\mathbf{QP}_{DNN}] fails to give a valid bound for instances 1, 3, 6, 7, 8 and 10, while the tensor relaxation [$\mathbf{TP} - \mathcal{K}^{DNN}$] can provide a valid lower bound for all tested instances.

6 Conclusion

This article presents convex relaxations for general POPs over CP and CPSD tensor cones. Bomze, in [8] shows that CP matrix relaxations are as tight as Lagrangian relaxations for QPs with both linear and quadratic constraints. A natural question is whether similar results hold for general POPs that are not necessarily quadratic. Introducing CP and CPSD tensors to reformulate or relax general POPs, we generalize Bomze's results to general POPs; that is, the CP tensor relaxations are as tight as Lagrangian relaxations for general POPs with degree higher than 2. These results provide another way of using symmetric tensor cones to globally approximate non-convex POPs. Burer in [13] shows that every QP with linear constraints and binary variables can be reformulated as a CP program and that QCQPs can be reformulated by CP programs under appropriate conditions. Note that one can reformulate general POPs as QPs by introducing additional variables and constraints and then apply Burer's results to obtain global bounds on general POPs. Peña et al. generalize Burer's results in [43] to show that under certain conditions a general POP can be reformulated as a conic program over the CP tensor cone. A natural question is which reformulations or relaxations will provide tighter bounds for general POPs. In this paper, we show that the bound of CP tensor relaxations is

tighter than the bound of CP matrix relaxations for the quadratic reformulation of some classes of general POPs. This validates the advantages of using tensor cones for convexification of non-convex POPs. We also provide some tractable approximations of the CP tensor cone as well as the CPSD tensor cone, which allows the possibility to compute the bounds based on these tensor relaxations. Some preliminary numerical results on small scale POPs show that these tensor cone approximations can provide good bounds for the global optimum of the original POPs. More importantly, in the numerical results performed, the bounds obtained by CP or CPSD tensor cone programs yield tighter bounds than the ones obtained with CP or SDP matrix relaxations for quadratic reformulation of general POPs using a similar computational effort. In the future, it will be interesting to further characterize the classes of POPs for which the CP and CPSD tensor cone relaxations provide tighter bounds than the CP and PSD matrix relaxations of its associated quadratic reformulations. Also, more POP instances with larger sizes can be tested and numerical comparisons on these more complicated POP cases can be made by developing appropriate code to address these problems.

Acknowledgements The authors would like to thank Sam Burer for his comments on an earlier version of this manuscript. Also, the authors would like to thank anonymous referees for their thorough and thoughtful comments. The authors of this article were supported by the NSF CMMI Grant 1300193.

References

- Ahmedi, A.A., Majumdar, A.: DSOS and SDSOS optimization: LP and SOCP-based alternatives to sum of squares optimization. In: Information Sciences and Systems (CISS), 2014 48th Annual Conference on, pp. 1–5. IEEE (2014)
- Ahmed, F., Dür, M., Still, G.: Copositive programming via semi-infinite optimization. *J. Optim. Theory Appl.* **159**(2), 322–340 (2013)
- Anjos, M.F., Lasserre, J.B.: *Handbook on Semidefinite, Conic and Polynomial Optimization*, vol. 166. Springer, Berlin (2011)
- Arima, N., Kim, S., Kojima, M.: A quadratically constrained quadratic optimization model for completely positive cone programming. *SIAM J. Optim.* **23**(4), 2320–2340 (2013)
- Arima, N., Kim, S., Kojima, M.: Extension of completely positive cone relaxation to moment cone relaxation for polynomial optimization. *J. Optim. Theory Appl.* **168**(3), 884–900 (2016)
- Bai, L., Mitchell, J.E., Pang, J.-S.: On conic QPCCs, conic QCQPs and completely positive programs. *Math. Program.* **159**(1–2), 109–136 (2016)
- Biswas, P., Lian, T.-C., Wang, T.-C., Ye, Y.: Semidefinite programming based algorithms for sensor network localization. *ACM Trans. Sens. Netw. (TOSN)* **2**(2), 188–220 (2006)
- Bomze, I.M.: Copositive relaxation beats Lagrangian dual bounds in quadratically and linearly constrained quadratic optimization problems. *SIAM J. Optim.* **25**(3), 1249–1275 (2015)
- Bomze, I.M., De Klerk, E.: Solving standard quadratic optimization problems via linear, semidefinite and copositive programming. *J. Glob. Optim.* **24**(2), 163–185 (2002)
- Bomze, I.M., Dür, M., De Klerk, E., Roos, C., Quist, A.J., Terlaky, T.: On copositive programming and standard quadratic optimization problems. *J. Glob. Optim.* **18**(4), 301–320 (2000)
- Bose, S., Low, S.H., Teeraratkul, T., Hassibi, B.: Equivalent relaxations of optimal power flow. *IEEE Trans. Autom. Control* **60**(3), 729–742 (2015)
- Bundfuss, S., Dür, M.: An adaptive linear approximation algorithm for copositive programs. *SIAM J. Optim.* **20**(1), 30–53 (2009)
- Burer, S.: On the copositive representation of binary and continuous nonconvex quadratic programs. *Math. Program.* **120**(2), 479–495 (2009)
- Burer, S., Anstreicher, K.M.: Second-order-cone constraints for extended trust-region subproblems. *SIAM J. Optim.* **23**(1), 432–451 (2013)
- Burer, S., Dong, H.: Representing quadratically constrained quadratic programs as generalized copositive programs. *Oper. Res. Lett.* **40**(3), 203–206 (2012)
- Cardoso, J.-F.: Blind signal separation: statistical principles. *Proc. IEEE* **86**(10), 2009–2025 (1998)
- Chen, J., Burer, S.: Globally solving nonconvex quadratic programming problems via completely positive programming. *Math. Program. Comput.* **4**(1), 33–52 (2012)

18. de Klerk, E., Pasechnik, D.V.: Approximation of the stability number of a graph via copositive programming. *SIAM J. Optim.* **12**(4), 875–892 (2002)
19. de Klerk, E., Pasechnik, D.V.: A linear programming reformulation of the standard quadratic optimization problem. *J. Glob. Optim.* **37**(1), 75–84 (2007)
20. Dickinson, P.J., Povh, J.: Moment approximations for set-semidefinite polynomials. *J. Optim. Theory Appl.* **159**(1), 57–68 (2013)
21. Dong, H.: Symmetric tensor approximation hierarchies for the completely positive cone. *SIAM J. Optim.* **23**(3), 1850–1866 (2013)
22. Dukanovic, I., Rendl, F.: Copositive programming motivated bounds on the stability and the chromatic numbers. *Math. Program.* **121**(2), 249–268 (2010)
23. Ghaddar, B., Marecek, J., Mevissen, M.: Optimal power flow as a polynomial optimization problem. *IEEE Trans. Power Syst.* **31**(1), 539–546 (2016)
24. Goemans, M.X.: Semidefinite programming in combinatorial optimization. *Math. Program.* **79**(1–3), 143–161 (1997)
25. Hu, S., Qi, L., Zhang, G.: Computing the geometric measure of entanglement of multipartite pure states by means of non-negative tensors. *Phys. Rev. A* **93**(1), 012304 (2016)
26. Jiang, B., Yang, F., Zhang, S.: Tensor and its tucker core: the invariance relationships. (2016) arXiv preprint [arXiv:1601.01469](https://arxiv.org/abs/1601.01469)
27. Kim, S., Kojima, M.: Exact solutions of some nonconvex quadratic optimization problems via SDP and SOCP relaxations. *Comput. Optim. Appl.* **26**(2), 143–154 (2003)
28. Kuang, X., Ghaddar, B., Naoum-Sawaya, J., Zuluaga, L.: Alternative LP and SOCP hierarchies for ACOFP problems. *IEEE Trans. Power Syst.* **32**, 2828–2836 (2016)
29. Lasserre, J.B.: Global optimization with polynomials and the problem of moments. *SIAM J. Optim.* **11**(3), 796–817 (2001)
30. Lasserre, J.B.: An explicit equivalent positive semidefinite program for nonlinear 0–1 programs. *SIAM J. Optim.* **12**(3), 756–769 (2002a)
31. Lasserre, J.B.: Semidefinite programming vs. LP relaxations for polynomial programming. *Math. Oper. Res.* **27**(2), 347–360 (2002b)
32. Lavaei, J., Low, S.H.: Zero duality gap in optimal power flow problem. *IEEE Trans. Power Syst.* **27**(1), 92–107 (2012)
33. Ling, C., Nie, J., Qi, L., Ye, Y.: Biquadratic optimization over unit spheres and semidefinite programming relaxations. *SIAM J. Optim.* **20**(3), 1286–1310 (2009)
34. Löfberg, J.: Yalmip: A toolbox for modeling and optimization in matlab. In: *IEEE International Symposium on Computer Aided Control Systems Design*, pp. 284–289. IEEE (2004)
35. Luo, Z., Qi, L., Ye, Y.: Linear operators and positive semidefiniteness of symmetric tensor spaces. *Sci. China Math.* **58**(1), 197–212 (2015)
36. Malek, A., Hosseini-pour-Mahani, N.: Solving a class of non-convex quadratic problems based on generalized KKT conditions and neurodynamic optimization technique. *Kybernetika* **51**(5), 890–908 (2015)
37. Mariere, B., Luo, Z.-Q., Davidson, T.N.: Blind constant modulus equalization via convex optimization. *IEEE Trans. Signal Process.* **51**(3), 805–818 (2003)
38. Mavridou, T., Pardalos, P., Pitsoulis, L., Resende, M.G.: A grasp for the biquadratic assignment problem. *Eur. J. Oper. Res.* **105**(3), 613–621 (1998)
39. Nesterov, Y.: Structure of non-negative polynomials and optimization problems. Technical report, Université catholique de Louvain, Center for Operations Research and Econometrics (CORE) (1997)
40. Parrilo, P.A.: Structured semidefinite programs and semialgebraic geometry methods in robustness and optimization. Ph.D. thesis, Citeseer (2000)
41. Peña, J., Vera, J., Zuluaga, L.F.: Computing the stability number of a graph via linear and semidefinite programming. *SIAM J. Optim.* **18**(1), 87–105 (2007)
42. Peña, J., Vera, J.C., Zuluaga, L.F.: A certificate of non-negativity for polynomials over unbounded sets. Lehigh University (2014a)
43. Peña, J., Vera, J.C., Zuluaga, L.F.: Completely positive reformulations for polynomial optimization. *Math. Program.* **151**(2), 405–431 (2014b)
44. Povh, J., Rendl, F.: A copositive programming approach to graph partitioning. *SIAM J. Optim.* **18**(1), 223–241 (2007)
45. Povh, J., Rendl, F.: Copositive and semidefinite relaxations of the quadratic assignment problem. *Discrete Optim.* **6**(3), 231–241 (2009)
46. QELA, R.E.B.E., KLINZ, B.: On the biquadratic assignment problem. In: *Quadratic Assignment and Related Problems: DIMACS Workshop, May 20–21, 1993*, vol. 16, p. 117. American Mathematical Soc (1994)

47. Shor, N.: Class of global minimum bounds of polynomial functions. *Cybern. Syst. Anal.* **23**(6), 731–734 (1987)
48. Song, Y., Qi, L.: Tensor complementarity problem and semi-positive tensors. *J. Optim. Theory Appl.* **169**, 1–10 (2015)
49. Vandenberghe, L., Boyd, S.: Semidefinite programming. *SIAM Rev.* **38**(1), 49–95 (1996)
50. Ye, Y.: Approximating quadratic programming with bound and quadratic constraints. *Math. Program.* **84**(2), 219–226 (1999)
51. Zheng, X.J., Sun, X.L., Li, D.: Convex relaxations for nonconvex quadratically constrained quadratic programming: matrix cone decomposition and polyhedral approximation. *Math. Program.* **129**(2), 301–329 (2011)
52. Zuluaga, L.F., Vera, J., Peña, J.: LMI approximations for cones of positive semidefinite forms. *SIAM J. Optim.* **16**(4), 1076–1091 (2006)